UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/764,557	01/27/2004	Andrew F. Knight		9594

42067 7590 10/04/2005

ANDREW F. KNIGHT
2770 AIRLINE GOLDMINE RD.
CANON, GA 30520

EXAMINER	
WANG, JIN CHENG	
ART UNIT	**PAPER NUMBER**
2672	

DATE MAILED: 10/04/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/764,557	KNIGHT, ANDREW F.	
	Examiner	Art Unit	
	Jin-Cheng Wang	2672	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *09 August 2005*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,2,4,5,7-13 and 15-23* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1,2,4,5,7-13 and 15-23* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendments

Applicant's submission on 8/9/2005 has been entered. Claims 1, 9, 12, and 18 have been amended. Claims 3, 6, and 14 have been canceled. Claims 21-23 have been newly added. Claims 1-2, 4-5, 7-13, 15-23 are pending in the present application.

Response to Arguments

Applicant's arguments filed on 8/9/2005 are moot in view of the new ground of rejection based on Cosatto et al. U.S. Patent No. 6,504,546. As set forth below in the Office Action, Cosatto discloses a method for creating a virtual video, comprising **at least one** of steps a)-f):

a) Sending an image of an object from a sender to a receiver via an information line (*e.g., the information line, be it wired or wireless, whether it is a physical line or a drawing path, is inherently associated with the system due to the image information exchange between the camera and the low-cost PC as a receiver or the information line is also inherently associated with the system due to the image information exchange between the database having a graphical interface module/processor for creating the bitmaps and the text-to-speech synthesizer/module/processor for receiving and then processing the bitmaps received from the database; column 12, lines 20-31 and column 15, lines 5-10. The text-to-audiovisual speech synthesizer processing audio and video streams is disclosed; column 14, lines 63-67*), said image having a plurality of identifiable image points, said plurality of identifiable image points (*e.g., samples for the bitmap of the facial parts are fewer than the remaining bitmap image points and the samples correspond to the object points of the facial object*) being substantially fewer in

number than a number of remaining image points of said image points of said image, said object having a plurality of identifiable object points, and said plurality of identifiable image points corresponding to said plurality of identifiable object points (*e.g., object refers to a three-dimensional object such as a talking person or the talking head of the base face; and the image refers to a bitmap of a facial part. In the case of modeling a human face, the set of three-dimensional planes correspond to a set of pre-defined facial parts and these bitmaps are then normalized and parameterized before being entered into a database. For the synthesis of a human head, a text-to-speech synthesizer provides the audio track, as well as a phoneme string and trajectory which computes motion for all the facial parts including the whole head. These trajectories provide the parameters for selecting the proper bitmaps from the database; see column 3, lines 27-53*);

b) repeatedly imaging the object to produce a first video (*Cosatto teaching recording a sequence of video and thus repeatedly imaging the object to produce a first video. Cosatto further discloses that, to create* ***a video animation at 30 frames per second and thereby repeatedly imaging the object to produce a first frame of video; see column 6, lines 50-67 and column 7, lines 1-13****, the trajectory is sampled every 33.33 milliseconds and for each sample point, the closest grid entry and its associated bitmap is chosen and the parameters describing feature shapes are chosen such that transitions between neighboring samples look smooth; column 13, lines 22-34; A frame of the final animation can be generated when bitmaps of all the face parts have been retrieved from the database and the bitmap of the base face is first copied into the frame buffer, then the bitmaps of face parts are projected onto the base face using the 3D model and the pose and the whole frame is rendered with just a few texture-map operations*

which makes it possible to render the talking head in real time on a low-cost PC; column 14, lines 62 to column 15, lines 9);

c) Determining, from the first video, object position data of said plurality of identifiable object points on said object (*e.g., Fig. 24 and Table 2 list various identifiable object points on the grid. Moreover, to create a video animation at 30 frames per second, the trajectory is sampled every 33.33 milliseconds and for each sample point, the closest grid entry and its associated bitmap is chosen and the parameters describing feature shapes are chosen such that transitions between neighboring samples look smooth; column 13, lines 22-34; A frame of the final animation can be generated when bitmaps of all the face parts have been retrieved from the database and the bitmap of the base face is first copied into the frame buffer, then the bitmaps of face parts are projected onto the base face using the 3D model and the pose and the whole frame is rendered with just a few texture-map operations which makes it possible to render the talking head in real time on a low-cost PC; column 14, lines 62 to column 15, lines 9*);

d) Sending said object position data to said receiver via an information line (*e.g., the receiver being a low-cost PC; column 15, lines 5-10 and an information line refers to the communication line between the low cost PC and the image bitmap database; column 14, lines 63-67*);

e) Morphing or warping said image such that image position data of said identifiable image points of said image are adjusted to approximately correspond to said object position data (*Morphing is discussed in the Background of Invention and the cited reference discloses that it is sufficient to use warping or alpha blending said image such that image position data of said identifiable image points of said image are adjusted to approximately correspond to said object*

position data for the purpose of computational saving; see column 13, lines 55-67. The cited reference teaches that morphing provides better results. During the transition interval, the resulting pixel is a blend of the corresponding pixels from sample a and sample b. The number of samples that are used to create a transition varies depending on the sampling rate of the trajectory and the duration of the samples. When the database contains few samples, the visual difference between samples is larger and more sophisticated techniques such as morphing provide better results. In column 14, the cited reference discloses that instead of directly mapping a phoneme to a viseme, each parameter of a viseme is derived from a sequence of phonemes and this generic model for coarticulation can be converted to a data-driven model and ***to synthesize new articulations of speech, the appropriate phoneme sequences*** *are identified in the coarticulation database and are then concatenated.*

Although the cited reference teaches warping or cheaper blending technique, it also teaches the claim limitation of "morphing" by disclosing the warping technique and the texture mapping technique for blending the image bitmaps and the base face model. The cited reference further discloses using morphing of the image bitmaps and the base face model to provide better results when the database contains few samples. Moreover, morphing has been extensively discussed in the Background of Invention. The cited reference teaches that morphing, warping and alpha blending for the texture mapping are the appropriate technique for smoothing and blending applied to the strings of bitmaps to eliminate hard transitions and ***create a seamless animation for each facial part*** *(column 3, lines 34-53 and Fig. 5, column 6, lines 7-20; column 7, lines 40-61). In column 7, lines 40-61, the cited reference further discloses a morphological*

operation followed by adaptive thresholding to result in a binary image where areas of facial features are marked with blobs of black pixels.

f) repeating steps c)-e) to produce a second video that substantially corresponds to the first video *(Cosatto discloses synthesizer that calculates motion trajectories for all of the facial parts as well as the base face wherein these trajectories provide the parameters for selecting the proper bitmaps from the database followed by the smoothing and blending to these strings of bitmaps to create a seamless animation for each facial parts and the talking head thus created* ***resembles very closely to the person who was original recorded,*** *i.e., the second video resembles very closely to the first video. Cosatto discloses recording real movements of a head and lips and reusing them for the synthesis to produce realistic lip and head movements as well as emotional expressions (column 3, lines 55-60). Moreover, Cosatto discloses the parameterization of the animation sequence describing the appearance of a facial part and thus the video sequence may be parameterized to generate another sequence. Cosatto further discloses that, to create a video animation at 30 frames per second; see column 6, lines 50-67 and column 7, lines 1-13 and column 10, lines 20-36 wherein the second video is substantially corresponds to the first video; see Fig. 26*), wherein all of steps a)-f) are performed.

Claim Rejections - 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claims 1-2, 4-5, 7-13, 15-23 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

A broad range or limitation together with a narrow range or limitation that falls within the broad range or limitation (in the same claim) is considered indefinite, since the resulting claim does not clearly set forth the metes and bounds of the patent protection desired. See MPEP § 2173.05(c). Note the explanation given by the Board of Patent Appeals and Interferences in *Ex parte Wu*, 10 USPQ2d 2031, 2033 (Bd. Pat. App. & Inter. 1989), as to where broad language is followed by "such as" and then narrow language. The Board stated that this can render a claim indefinite by raising a question or doubt as to whether the feature introduced by such language is (a) merely exemplary of the remainder of the claim, and therefore not required, or (b) a required feature of the claims. Note also, for example, the decisions of *Ex parte Steigewald*, 131 USPQ 74 (Bd. App. 1961); *Ex parte Hall*, 83 USPQ 38 (Bd. App. 1948); and *Ex parte Hasche*, 86 USPQ 481 (Bd. App. 1949).

In the present instance, claim 1 recites the broad recitation "comprising at least one of steps (a)-f)", and the claim also recites "wherein all of steps a)-f) are performed" which is the narrower statement of the range/limitation. Claims 2, 4-5, 7-11 and 21 depend upon the claim 1 and are rejected due to their dependency on the claim 1.

Claim 12 recites the broad recitation "comprising at least one of steps (a)-h)", and the claim also recites "wherein all of steps a)-h) are performed" which is the narrower statement of the range/limitation. The claims 13, 15-17 and 22 depend upon the claim 12 and are rejected due to their dependency on the claim 12.

Claim 18 recites the broad recitation "comprising at least one of steps (a)-h)", and the claim also recites "wherein all of steps a)-h) are performed" which is the narrower statement of the range/limitation. The claims 19-20 and 23 depend upon the claim 18 and are rejected due to their dependency on the claim 18.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

Claims 1-2, 4-5, 7-9, 12-13, 15-20 are rejected under 35 U.S.C. 102(e) as being anticipated by Cosatto et al. U.S. Patent No. 6,504,546 (hereinafter Cosatto).

Re Claims 1, 12, 18:

Cosatto discloses a method for creating a virtual video, comprising at least one of steps a)-d):

a) Sending an image of an object from a sender to a receiver via an information line (*e.g., the information line, be it wired or wireless, whether it is a physical line or a drawing path, is inherently associated with the system due to the image information exchange between the camera and the low-cost PC as a receiver or the information line is also inherently associated with the system due to the image information exchange between the database having a graphical*

interface module/processor for creating the bitmaps and the text-to-speech synthesizer/module/processor for receiving and then processing the bitmaps received from the database; column 12, lines 20-31 and column 15, lines 5-10. The text-to-audiovisual speech synthesizer processing audio and video streams is disclosed; column 14, lines 63-67), said image having a plurality of identifiable image points, said plurality of identifiable image points (*e.g., samples for the bitmap of the facial parts are fewer than the remaining bitmap image points and the samples correspond to the object points of the facial object*) being substantially fewer in number than a number of remaining image points of said image points of said image, said object having a plurality of identifiable object points, and said plurality of identifiable image points corresponding to said plurality of identifiable object points (*e.g., object refers to a three-dimensional object such as a talking person or the talking head of the base face; and the image refers to a bitmap of a facial part. In the case of modeling a human face, the set of three-dimensional planes correspond to a set of pre-defined facial parts and these bitmaps are then normalized and parameterized before being entered into a database. For the synthesis of a human head, a text-to-speech synthesizer provides the audio track, as well as a phoneme string and trajectory which computes motion for all the facial parts including the whole head. These trajectories provide the parameters for selecting the proper bitmaps from the database; see column 3, lines 27-53*);

b) Repeatedly imaging the object to produce a first video (*Cosatto discloses in column 1, lines 25-30 recording video clips of real people or cartoon characters and recording real movements of a head and lips and reusing them for the synthesis to produce realistic lip and head movements as well as emotional expressions;* see column 3, lines 55-60. *Cosatto teaching*

recording a sequence of video and thus repeatedly imaging the object to produce a first video. Cosatto further discloses that, to create ***a video animation at 30 frames per second and thereby repeatedly imaging the object to produce a first frame of video; see column 6, lines 50-67 and column 7, lines 1-13****, the trajectory is sampled every 33.33 milliseconds and for each sample point, the closest grid entry and its associated bitmap is chosen and the parameters describing feature shapes are chosen such that transitions between neighboring samples look smooth; column 13, lines 22-34; A frame of the final animation can be generated when bitmaps of all the face parts have been retrieved from the database and the bitmap of the base face is first copied into the frame buffer, then the bitmaps of face parts are projected onto the base face using the 3D model and the pose and the whole frame is rendered with just a few texture-map operations which makes it possible to render the talking head in real time on a low-cost PC; column 14, lines 62 to column 15, lines 9*);

c) Determining, from the first video, object position data of said plurality of identifiable object points on said object (*e.g., Fig. 24 and Table 2 list various identifiable object points on the grid. Moreover, to create a video animation at 30 frames per second, the trajectory is sampled every 33.33 milliseconds and for each sample point, the closest grid entry and its associated bitmap is chosen and the parameters describing feature shapes are chosen such that transitions between neighboring samples look smooth; column 13, lines 22-34; A frame of the final animation can be generated when bitmaps of all the face parts have been retrieved from the database and the bitmap of the base face is first copied into the frame buffer, then the bitmaps of face parts are projected onto the base face using the 3D model and the pose and the whole frame*

is rendered with just a few texture-map operations which makes it possible to render the talking head in real time on a low-cost PC; column 14, lines 62 to column 15, lines 9);

d) Sending said object position data to said receiver via an information line (*e.g., the receiver being a low-cost PC; column 15, lines 5-10 and an information line refers to the communication line between the low cost PC and the image bitmap database; column 14, lines 63-67*);

e) Morphing or warping said image such that image position data of said identifiable image points of said image are adjusted to approximately correspond to said object position data (*Morphing is discussed in the Background of Invention and the cited reference discloses that it is sufficient to use warping or alpha blending said image such that image position data of said identifiable image points of said image are adjusted to approximately correspond to said object position data for the purpose of computational saving; see column 13, lines 55-67. The cited reference teaches that morphing provides better results. During the transition interval, the resulting pixel is a blend of the corresponding pixels from sample a and sample b. The number of samples that are used to create a transition varies depending on the sampling rate of the trajectory and the duration of the samples. When the database contains few samples, the visual difference between samples is larger and more sophisticated techniques such as morphing provide better results. In column 14, the cited reference discloses that instead of directly mapping a phoneme to a viseme, each parameter of a viseme is derived from a sequence of phonemes and this generic model for coarticulation can be converted to a data-driven model and* ***to synthesize new articulations of speech, the appropriate phoneme sequences*** *are identified in the coarticulation database and are then concatenated.*

Although the cited reference teaches warping or cheaper blending technique, it also teaches the claim limitation of "morphing" by disclosing the warping technique and the texture mapping technique for blending the image bitmaps and the base face model. The cited reference further discloses using morphing of the image bitmaps and the base face model to provide better results when the database contains few samples. Moreover, morphing has been extensively discussed in the Background of Invention. The cited reference teaches that morphing, warping and alpha blending for the texture mapping are the appropriate technique for smoothing and blending applied to the strings of bitmaps to eliminate hard transitions and ***create a seamless animation for each facial part*** *(column 3, lines 34-53 and Fig. 5, column 6, lines 7-20; column 7, lines 40-61). In column 7, lines 40-61, the cited reference further discloses a morphological operation followed by adaptive thresholding to result in a binary image where areas of facial features are marked with blobs of black pixels.*

f) Repeating steps c)-e) to produce a second video that substantially corresponds to the first video *(Cosatto discloses synthesizer that calculates motion trajectories for all of the facial parts as well as the base face wherein these trajectories provide the parameters for selecting the proper bitmaps from the database followed by the smoothing and blending to these strings of bitmaps to create a seamless animation for each facial parts and the talking head thus created* ***resembles very closely to the person who was original recorded****, i.e., the second video resembles very closely to the first video. Cosatto discloses recording real movements of a head and lips and reusing them for the synthesis to produce realistic lip and head movements as well as emotional expressions (column 3, lines 55-60). Moreover, Cosatto discloses the parameterization of the animation sequence describing the appearance of a facial part and thus the video sequence may*

be parameterized to generate another sequence. Cosatto further discloses that, to create a video animation at 30 frames per second; see column 6, lines 50-67 and column 7, lines 1-13 and column 10, lines 20-36 wherein the second video is substantially corresponds to the first video; see Fig. 26), wherein all of steps a)-f) are performed.

The claims 12 and 18 further recite "recording voice information of a human", however, Cosatto discloses in column 4, lines 10-25 capturing accurately realistic speech postures and human subjects speak short text sequences in front of a camera and a face recognition system then automatically analyzes this video footage and selects the proper samples. A sequence of phonemes are captured and therefore Cosatto teaches the claim limitation.

With regards to the sender processor and receiver processor as recited in the claims 12 and 18, it is known that Cosatto discloses processing modules associated with the sender, i.e., the camera or the database having the associated user interface modules, and the receiver, the PC having the microprocessor or having the synthesizer module for synthesizing the virtual video such as the live talking head.

Claim 2:

Cosatto further discloses morphing said image bitmaps such that image position data of said remaining image points are adjusted depending on said object position data (column 7, lines 50-61 and column 11, column 14, line 62 to column 15, lines 9).

Claim 4:

Cosatto further discloses three-dimensional object position data of a talking head (Fig. 24 and Table 2 and column 11).

Claim 5:

Cosatto further discloses the animation of the remaining facial parts including jaw, eyes, forehead and eyebrows and identifying and determining the remaining facial parts include identifying and determining the second identifiable image points corresponding to the second identifiable image points of the base face (column 14, lines 53-61).

Claim 7:

Cosatto further discloses in column 14, lines 62-67 that a frame of the final animation can be generated when bitmaps of all the face part have been retrieved from the database and the bitmap of the base face is first copied into the frame buffer and then the bitmaps of face parts are projected onto the base face using the 3D model and the pose. The second image and the third image refer to the second bitmap and the third bitmap of the facial parts. The first frame and the second refer to the first frame and the second in a sequence of viseme. With regards to the identifiable image points, Cosatto discloses in column 10, lines 50-53 that the outline of lips, one of the facial parts, for example, encoded as a sequence of points and all these points are then mapped into the normalized plane before entering them into the database. With regards to the object points, Cosatto further discloses in Fig. 24 and Table 2 a list of various identifiable object points on the grid. With regards to the relationship between the first frame and the second, **Cosatto discloses that, to create a video animation at 30 frames per second**, the trajectory is sampled every 33.33 milliseconds and for each sample point, the closest grid entry and its

associated bitmap is chosen and the parameters describing feature shapes are chosen such that transitions between neighboring samples look smooth; column 13, lines 22-34; A frame of the final animation can be generated when bitmaps of all the face parts have been retrieved from the database and the bitmap of the base face is first copied into the frame buffer, then the bitmaps of face parts are projected onto the base face using the 3D model and the pose and the whole frame is rendered with just a few texture-map operations which makes it possible to render the talking head in real time on a low-cost PC; column 14, lines 62 to column 15, lines 9.

Claim 8:

Cosatto further discloses that the number of samples that are used to crate a transition varies depending on the sampling rate of the trajectory and the duration of the samples (column 13, lines 55-67).

Claim 9:

Cosatto further discloses capturing tens of thousands of video frames (column 15, lines 30-46), training a set of 300 frames (column 8, lines 50-55) and using a variety of frame rates including a rate of at least 5 times per second (column 6, line 66 to column 7, line 13).

Re Claim 13:

Cosatto further discloses viewing a facial image as viseme (column 15, lines 10-19) and marking an area by the color analysis as a candidate of a face area combined with candidates of eye areas produced by the texture analysis (column 7, lines 50-61) and marking on the shape of the lips of the current phoneme being uttered (column 14) and mapping a phoneme to a viseme (column 14).

Re Claims 15 and 20:

Cosatto discloses capturing accurately realistic speech postures, human subjects speaking short text sequences in front of a camera and automatically analyzing the video footage by the face recognition system and selecting the proper samples and extracting the needed bitmaps from video frames and synthesizing the talking head animation to create the photo-realistic talking head (column 4, lines 10-22). Cosatto discloses mapping a phoneme to a viseme (column 14) and using the text-to-speech synthesizer to drive the entire animation to create a talking head (column 15). Cosatto further discloses that morphing, warping and alpha blending for the texture mapping are the appropriate technique for smoothing and blending applied to the strings of bitmaps to eliminate hard transitions and create a seamless animation for each facial part (column 3, lines 34-53 and Fig. 5, column 6, lines 7-20; column 7, lines 40-61). In column 7, lines 40-61, the cited reference further discloses a morphological operation followed by adaptive thresholding to result in a binary image where areas of facial features are marked with blobs of black pixels.

Re Claim 16:

Cosatto further discloses morphing the remaining facial parts such as jaw, eyes, forehead and eyebrows (column 14).

Re Claim 17:

Cosatto further discloses capturing tens of thousands of video frames (column 15, lines 30-46), training a set of 300 frames (column 8, lines 50-55) and using a variety of frame rates including a rate of at least 5 times per second (column 6, line 66 to column 7, line 13). Cosatto discloses displaying a virtual video of a talking head (column 15).

Re Claim 19:

Cosatto discloses high-resolution animation involving the short sequences for the base face totaling about 3MB compressed using MPEG 2 and the facial parts including jaw, eyes, forehead and eyebrows of 5 kB for each sample with a total of 40 samples and 48 mouth samples to create the sound face image (column 11, line 39 to column 12, line 5).

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claim 10 is rejected under 35 U.S.C. 103(a) as being unpatentable over Cosatto et al. U.S. Patent No. 6,504,546 (hereinafter Cosatto) in view of Hayashi U.S. Patent No. 5,652,670 (hereinafter Hayashi).

Cosatto further discloses recording a person's posture using cameras (column 6, lines 50-65) and using the 3D scanning techniques such as a CyberWare range scanner (column 1, lines 50-65). Cosatto is silent to using the laser based scanners and cameras. However, Hayashi discloses a laser scanner (See Hayashi the Abstract). It would have been obvious to have used Hayashi's laser scanner for taking a person's facial image because Cosatto has taught using a CyberWare range scanner or an optical scanner (column 1, lines 50-65) which may be a laser scanner by itself, or if not, alternatively using Hayashi's laser scanner because at the time of invention, a laser scanner is available for taking a person's facial image. One of the ordinary skill

in the art would have been motivated do incorporate an optical scanner such as a laser scanner for taking a person's facial image using a compact scanner for cost reduction (Hayashi column 1).

Claim 11 is rejected under 35 U.S.C. 103(a) as being unpatentable over Cosatto et al. U.S. Patent No. 6,504,546 (hereinafter Cosatto).

Cosatto further discloses machine-executable code (Table 3 and column 12, lines 53-57) to cause a machine (PC) to perform the method as in claim 1. Cosatto however is silent to a computer-usable medium. However, one of ordinary skill in the art would have recognized that computer usable medium (i.e., floppy, cd-rom, etc.) carrying computer-executable instructions for implementing a method, because it would facilitate the transporting and installing of the method on other systems, is generally well-known in the art. For example, a copy of the Microsoft Windows operating system can be found on a cd-rom from which Windows can be installed onto other systems, which is a lot easier than running a long cable or hand typing the software onto another system. The Office takes Official Notice of this teaching. Therefore, it would have been obvious to put Cosatto's program or algorithm on a computer readable medium, because it would facilitate the transporting, installing and implementing of Cosatto's program or algorithm on other systems.

Claims 21-23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Cosatto et al. U.S. Patent No. 6,504,546 (hereinafter Cosatto).

Cosatto discloses e.g., the information line is inherently associated with the system due to the image information exchange between the camera and the low-cost PC as a receiver or the

information line is also inherently associated with the system due to the image information exchange between the database having a graphical interface processor for creating the bitmaps and the text-to-speech synthesizer/processor for receiving and then processing the bitmaps received from the database; column 12, lines 20-31 and column 15, lines 5-10. The text-to-audiovisual speech synthesizer processing audio and video streams is disclosed; column 14, lines 63-67. The claims 21-23 recite the sender and the receiver being located in different cities. However, it does not matter the database server is located closely or remotely or in different cities from the PC processing the animation functions such as a talking head because at the time of the invention was made, the internet cable or dial-up is readily prevalent and the communication lines among computers of different cities are readily available to the general public. The Office takes Official Notice of this teaching. Therefore, it would have been obvious to construct Cosatto's method such that the database or the camera is being located remotely in a city different from where the PC is located so that the talking head or video clips are being executed for a person located remotely in a different city near a camera or for the person remotely located with the pre-recorded video in a database wherein the video clips are sent over the communication lines available to the general public.

Conclusion

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jin-Cheng Wang whose telephone number is (571) 272-7665. The examiner can normally be reached on 8:00 - 6:30 (Mon-Thu).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Mike Razavi can be reached on (571) 272-7664. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

jcw

MICHAEL RAZAVI
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600